THE PHYSICS OF B FACTORIES

P.F. Harrison
Physics Department, Queen Mary and Westfield College Mile End Rd. London E1 4NS. UK

Abstract
This course of two lectures reviews the physics goals of B factories, and compares and contrasts the different types of B factories. It then reviews the physics of CP violation in B decays, the ways of observing it, and the expected sensitivity to the fundamental quantities.

1. INTRODUCTION TO B FACTORIES

1.1 What is a B factory?

For the purposes of these lectures, I take as my definition 'any facility, built specifically to do B physics, which produces $> 10^7$ B mesons per year'. There are several types of B factory:

- e^+e^- symmetric, eg. the CESR storage ring at Cornell and the CLEO experiment which is located there.
- e^+e^- asymmetric, eg. the PEP-II collider and the associated BABAR experiment at SLAC; the KEK-B machine and the BELLE experiment at KEK.
- Fixed target hadronic, eg. the HERA-B experiment at the HERA machine at DESY.
- Hadron Collider, eg. the LHC-B experiment at CERN and the B-TEV experiment at Fermilab.

1.2 Physics Goals of B Factories

To do B physics! We can categorise the goals of B factories in terms of the kind of machine which is needed to achieve them. The goals which need fast-moving B's can only be done at e^+e^- asymmetric B factories, *and* at hadronic machines. These goals are those which rely on observations of the time-dependence of B decays, namely

- B lifetime measurements
- Time-dependent $B^0 - \overline{B}^0$ mixing
- Time-dependent CP-violating asymmetries

Each of the above is related to measurements of parameters of the CKM matrix, introduced in Daniel Wyler's lectures [1], and reviewed in Section 2.1. The B lifetime is one component needed in the determination of the CKM parameter $|V_{cb}|$, $B^0 - \overline{B}^0$ mixing gives information on $|V_{td}|$ and $|V_{ts}|$, while time-dependent CP-violating asymmetries measure the CP angles α and β of the unitarity triangle [1], shown in Fig. 1. In particular, time-dependent CP-violating asymmetries in B decays to low-multiplicity final states are useful, as follows:

- Channels useful for measuring sin(2β)
  - $B^0 \to J/\psi\, K^0_S, \quad B^0 \to J/\psi\, K^0_L, \quad B^0 \to J/\psi\, K^*$
  - $B^0 \to D^+D^-, \quad B^0 \to D^{*+}D^-, \quad B^0 \to D^{*+}D^{*-}$
- Channels useful for measuring α
  - $B^0 \to \pi^+\pi^-$
  - $B^0 \to (\rho^+\pi^- \text{or } \rho^-\pi^+) \to \pi^+\pi^-\pi^0$
  - $B^0 \to \rho\rho, \quad B^0 \to a_1\pi$

Time-dependent asymmetries are needed because much (at hadronic mahines) or all (at e^+e^- machines) of the asymmetries cancel in time-integrated measurements. Of course, these CP violation measurements are the main *raisons d'etre* of BABAR, BELLE, and hadronic B factories, and the experiments'

success will be judged by how well they perform in this area. We will see in detail later how these obsevations lead to the determinations of the CP angles concerned, and also consider how the observations are made, in practice.

The remaining physics goals for B factories can be made at all of the B factories, symmetric or asymmetric e^+e^-, or hadronic machines:

- Rare B decays: many of the channels cited in the section above, and the analogous decay modes of charged B mesons, are presently unobserved and can be considered themselves as rare decay modes. The branching ratios are interesting in their own right, and are being measured. In addition, the following rare decay modes are interesting:
  - $B^0 \to K^\pm\pi^\mp$, $B^0 \to K\rho$, $B^0 \to K^*\pi$ etc. to measure the relative size of penguin decays, and also perhaps to observe direct CP-violation
  - $B \to \eta h$, $B \to \eta' h$, where h is π or K, which may be interesting for direct CP-violation studies
  - Searches for non-Standard Model physics in $b \to s\gamma$, $b \to d\gamma$, $b \to sl^+l^-$, $b \to s\nu\nu$, $b \to s\gamma\gamma$, $b \to sg$
  - Leptonic B decays
- Improved determintions of V_{cb} using:
  - exclusive $B \to D^*(D)l\nu$ decays. Heavy Quark Effective Theory enables the extraction of V_{cb} from this channel with competitive errors, at the same time as determining various form factors
  - inclusive semileptonic decays
- Improved determintions of V_{ub} using:
  - inclusive decays, eg. by the lepton spectrum endpoint method
  - exclusive decays such as $B \to \rho l\nu$, $B \to \pi l\nu$
- Direct CP-violation using $B \to DK$ rate asymmetries
- Charm Physics
- τ Physics

1.3 e^+e^- B Factories

e^+e^- B factories operate at the $\Upsilon(4\mathrm{S})$ resonance. Fig. 2 shows the first four Υ resonances, as a function of centre-of-mass energy in e^+e^- collisions (the data are from the CUSB experiment at CESR). Although the $\Upsilon(4\mathrm{S})$ has a smaller cross-section than the first three resonances, it is used because it is the first one which is massive enough to produce B mesons in its decay. The cross-sections for different types of final state in e^+e^- collisions at the $\Upsilon(4\mathrm{S})$ resonance are shown in Table 1. It is easy to see that the ratio of

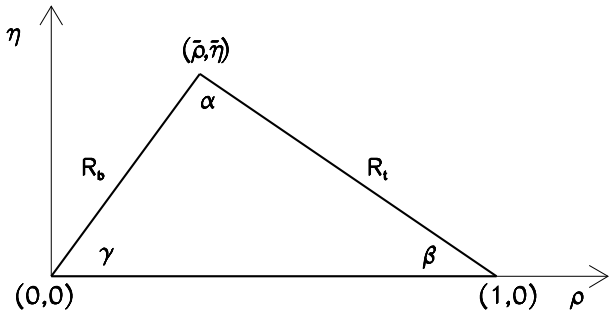

Fig. 1: The unitarity triangle, which parameterises two of the degrees of freedom of the CKM matrix, and whose angles may be determined by measuring time-dependent CP asymmetries at B factories.

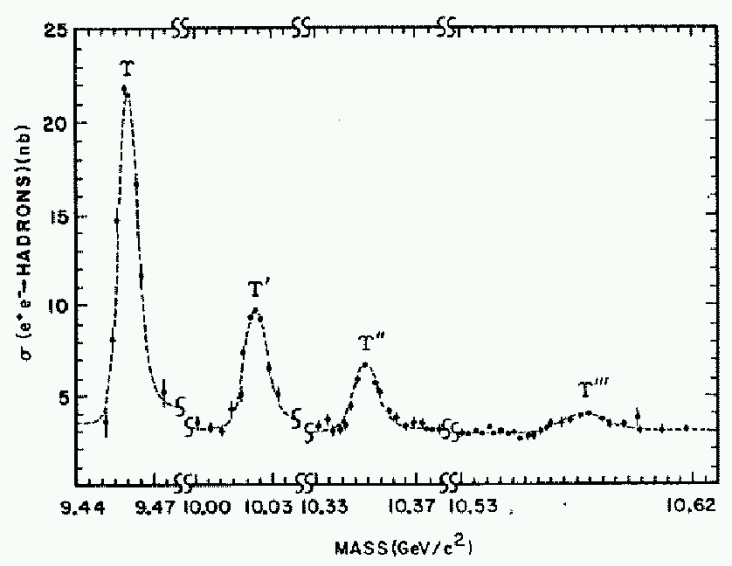

Fig. 2: The cross-section for $e^+e^- \to$ hadrons in the range $\sqrt{s} =$ 9.44-10.62 GeV at the CUSB experiment at CESR, showing the first four Υ resonances.

the $b\overline{b}$ cross-section to the total hadronic cross-section is ~ 0.22 at this energy. This is higher than the B fraction at LEP, and orders of magnitude higher than can be obtained in hadronic collisions, explaining why the $\Upsilon(4\mathrm{S})$ resonance is typically the energy of choice for e^+e^- B factories.

The $\Upsilon(4\mathrm{S})$ resonance is only fractionally above threshold for B mesons, so that, when produced, the B mesons are moving very slowly in the centre of mass system (cms) of the $\Upsilon(4\mathrm{S})$, Fig. 3. You can calculate as an exercise that the momentum of the B is only about 340 MeV/c, and that its relativistic velocity, β is ~ 0.06 in the cms. In other words, the B's are almost at rest (in relativistic terms) in the cms, and only travel of order $\sim 30\mu$m before decaying. This explains why, in symmetric e^+e^- collisions at the $\Upsilon(4\mathrm{S})$, it is not possible to measure their lifetime, or other time-dependent features of their decays: most vertex detector systems have a spatial resolution larger than this. At hadronic machines (and even at LEP), the produced B mesons are highly relativistic, making lifetime-dependent measurements possible. At the $\Upsilon(4\mathrm{S})$, we need to engineer asymmetric collisions, so that the B mesons are moving fast enough to travel a measurable distance before they decay.

Table 1: Production cross-sections at $\sqrt{s} = M(\Upsilon(4\mathrm{S}))$. The e^+e^- cross-section is the effective cross-section, expected within the experimental acceptance.

$e^+e^- \to$	Cross-section (nb)
$b\overline{b}$	1.05
$c\overline{c}$	1.30
$s\overline{s}$	0.35
$u\overline{u}$	1.39
$d\overline{d}$	0.35
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

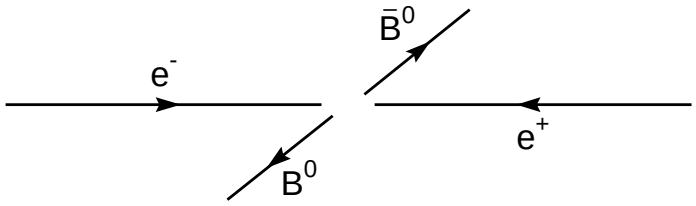

Fig. 3: The two B mesons are produced at non-relativistic velocities in symmetric collisions at the $\Upsilon(4S)$.

Fig. 4: The two B mesons are produced at relativistic velocities, with $\beta\gamma \sim 0.5$ in asymmetric collisions at the $\Upsilon(4S)$, giving measurable decay lengths. The B^0 momenta are close to the collision axis and are nearly equal to each other.

At asymmetric e^+e^- B factories, the electron and postitron are made to collide head-on, but with different momenta in the cms, as shown in Fig. 4. In the final state, the resulting B mesons each have a momentum $p_B \sim \frac{|p(e^-)-p(e^+)|}{2}$, which is typically large compared with their cms momenta (the cms momentum still characterises the magnitude of their transverse momentum in the lab. frame). You can again satisfy yourself that the mean flight distance of a B meson in the lab is $\Delta z = \beta\gamma c\tau_B$ where the relativistic factor $\beta\gamma = p_B/M_B$. Table 2 shows these various factors for the two B factories PEP-II and KEK-B.

Table 2: Summary of asymmetric e^+e^- B factory kinematics.

Quantity	PEP-II/BABAR	KEK-B/BELLE
$p(e^-)$ (GeV/c)	9	8
$p(e^+)$ (GeV/c)	3.1	3.5
$< p_B >$ (GeV/c)	2.95	2.25
$\beta\gamma$	0.56	0.43
$< \Delta z >$ (μm)	260	200

The design luminosities of the two asymmetric machines are $3(10) \times 10^{33}$ cm^{-2}s^{-1} for PEP-II (BELLE) respectively, although, at the time of writing, PEP-II has delivered more luminosity than KEK-B, and it is anticipated that PEP-II should also be able to reach the higher of these two goals, eventually. Given a cross-section of just over 1 nb for $b\overline{b}$ production, the number of $B\overline{B}$ pairs is expected to be $3(10) \times 10^7$ per year at BABAR (BELLE) (a particle physics year is usually estimated at 10^7 seconds, about a third of its actual length, owing to the fact that particle accelerators and detectors cannot be run continuously, and need time out for maintenance). These are divided equally between $B^0\overline{B}^0$ pairs and B^+B^- pairs.

There are several advantages of the e^+e^- environment over the hadronic environment, namely:

- It is possible to trigger on, and record, essentially all $B\overline{B}$ events.
- The high signal-to-background ratio, with $\sigma_{b\overline{b}}/\sigma_{hadr} \simeq 0.22$.
- The events are relatively clean, with a mean charged multiplicty of ~ 11.
- In B events, the B and $\overline{B}$ particles are the only particles in the final state.
- The interaction rates are low, ~ 10 Hz (physics rate).
- The possibility to reconstruct final states containing π^0s and photons, thereby allowing the possibility to make measurements in many more channels.
- Straightforward extrapolation from existing experiments.

Disadvantages of the e^+e^- environment compared with the hadronic environment are

- Statistics are limited to a few $\times 10^7$ $B\overline{B}$ pairs per year.
- The B_s^0 meson is not created at the $\Upsilon(4\mathrm{S})$.

It is possible to run e^+e^- B factories at the $\Upsilon(5\mathrm{S})$, but the cross-section is much reduced there, and it is not presently considered that this will provide a competitive advantage over other techniques of reaching these mesons, such as hadronic collisions.

1.4 Hadronic B factories

These can further be categorised into fixed-target and collider, although the distinctions are not as great as might initially be assumed.

1.4.1 Fixed Target Hadronic B Factories

The HERA-B experiment at DESY is the obvious example of the fixed-target hadronic B factory. It uses a wire target in the halo of the proton beam at HERA, providing collisions between protons and nucleons at $\sqrt{s} \sim 43$ GeV. The B mesons are emitted in the forward direction with a typical energy of order 100 GeV. The collisions provide an effective design luminosity of 3.5×10^{33} cm^{-2}s^{-1}, which, with a cross-section ~ 12 nb, should provide about 4×10^8 $B\overline{B}$ pairs per year. Although the production rate of B mesons is greater in a hadronic B factory, the environment is much more demanding, and the challenge is to dig the signal out from the very large flux of non-B events. After triggering and reconstruction, the total number of useful events remaining is cut dramatically, as is the range of available final states for study. The trigger of HERA-B requires high tranverse momentum leptons and/or a combination of 2 tracks with $\simeq$ a J/ψ mass.

Advantages of the fixed-target environment over the e^+e^- environment, include:

- A long B flight length, of order 1 cm.
- The availability of B_s^0 mesons for study.
- A simple (planar) detector geometry.

Disadvantages of the fixed-target environment include:

- The B decay products are highly collimated
- $\sigma_{b\overline{b}}/\sigma_{tot} \simeq 10^{-6}$.

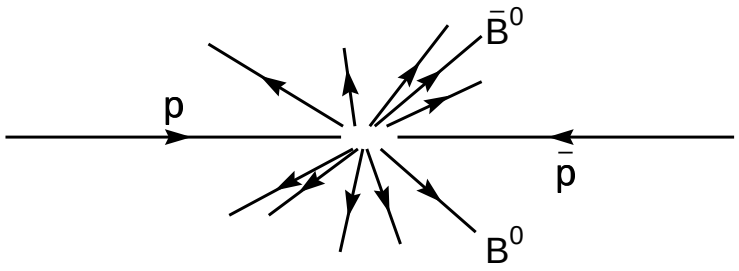

Fig. 5: The two B mesons are produced strongly forward-peaked and on the same side of the events in symmetric hadronic collisions.

- The mean charged multiplicity is $\simeq 60$.
- The interacton rate is ~ 40 MHz (with 4 interactions per beam crossing of the target).
- the detector occupancies are huge
- there are fake particle/anti-particle asymmetries in the production rates of B and $\overline{B}$.

1.4.2 Hadron Collider B Factories

Two hadron collider B factory experiments have been proposed. The B-TEV experiment at Fermilab's Tevatron collider and the LHC-B experiment at CERN's LHC. In these experiments, protons will be brought to collide head-on at high energy ($\sqrt{s} \sim 2$ GeV, 14 GeV respectively) producing $B\overline{B}$ pairs copiously in the forward direction. Somewhat counter-intuitively, both the B and $\overline{B}$ are found in simulations to be produced on the same side of the collisions, as shown in Fig. 5. This is explained by the fact that they are produced by hard parton-parton collisions, and it is extremely unlikely that both partons in a given collision will contain equal and opposite momenta. In fact, the situation is comparable to an asymmetric parton collider! The $B\overline{B}$ pair has a momentum equal to the vector sum of the two partons' momenta, p_1 and p_2, so that each B meson has a momentum magnitude approximately equal to $|p_1 - p_2|/2$, in analogy with the e^+e^- asymmetric collider case. This turns out, typically, to be in the range 1-100 GeV. With luminosities in the range of a few $\times 10^{32}$ cm^{-2}s^{-1} and cross-sections of a few $\times 10^{-28}$ cm^2, several $\times 10^{11}$ $B\overline{B}$ pairs are expected to be produced annually. With a high p_T trigger, these experiments are expected to have access to the full range of physics goals listed in Section 1.2, althouth it is likely that few channels with π^0s and γs in the final state will be observable, given the huge average multiplicities of the events.

Advantages of the hadron collider environment include:

- The availability of B_s^0 mesons for study.
- A simple (planar) detector geometry.
- huge $b\overline{b}$ cross-sections.
- $\sigma_{b\overline{b}}/\sigma_{tot} \sim$ a few $\times 10^{-3}$

Disadvantages of the fixed-target environment include:

- High multiplicity events.
- Difficult triggering.
- High interaction rate ($\sim$ 20 MHz).
- There are fake particle/anti-particle asymmetries in the production rates of B and $\overline{B}$.

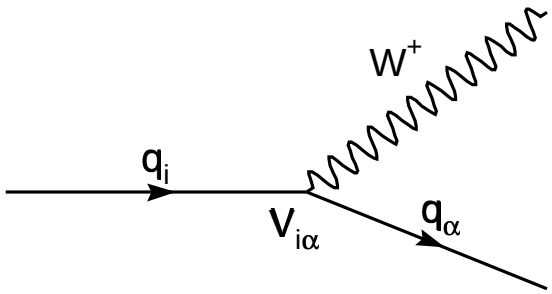

Fig. 6: Charged-current W vertex, whose strength depends on the types of quarks on either side and is proportional to the CKM matrix element $V_{i\alpha}$, where the charge 2/3 quark has flavour i and the charge -1/3 quark has flavour α.

1.5 Summary

Table 3 summarises several of the relevant parameters which characterise the various kinds of B factories described in this section.

Table 3: Some of the relevant parameters which charactoerise e^+e^- and hadronic B factories.

	e^+e^- B factory	HERA-B	B-TEV	LHC-B
$\sqrt{s}$	10.6	43	2×10^3	1.4×10^4
$b\overline{b}$ Cross-sect.	1.1×10^{-33}	1.2×10^{-32}	1.0×10^{-28}	1.0×10^{-28}
Luminosity ($cm^{-2}s^{-1}$)	$(3-10) \times 10^{33}$	3.5×10^{33}	2.0×10^{32}	2.0×10^{32}
No. $b\overline{b}$ yr^{-1}	$(3-10) \times 10^7$	$4.0 \times 10^8 \times \epsilon$	$2 \times 10^{11} \times \epsilon$	$1 \times 10^{12} \times \epsilon$
$\sigma_{b\overline{b}}/\sigma_{tot}$	0.22	1.0×10^{-6}	2.0×10^{-3}	6.0×10^{-3}

2. B FACTORY PHYSICS

There are many excellent texts on this area [2, 3, 4, 5]. Perhaps the most exhaustive is the BABAR Physics Book [2], which itself contains over 2000 references on the subject. I have therefore refrained from a large bilbliography, preferring instead to refer to just the short list above.

2.1 The CKM matrix and B physics

The CKM matrix parameterises the couplings of the quarks to the $W^{\pm}$ as shown in Fig. 6. With three generations of quarks, there are nine combinations which can participate in such interactions, and their relative couplings are conventionally represented as the elements of the 3×3 mixing matrix which is shown in Eq. (1). The matrix, however, is not simply an empirical parameterisation of coupling constants. It arises naturally from the gauge theory of weak interactions, and its form is constrained by that theory. In particular, universality of the weak interactions (which is required in a gauge theory) requires that the mixing matrix be unitary. Choosing appropriate phases for the quark fields, such a matrix may be parameterised using only four parameters (3 magnitudes and one phase) as has been done elegantly by Wolfenstein [2]:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\lambda^4, \qquad (1)$$

where $\lambda \simeq 0.22$ and A, ρ and η are of order unity. The unitarity of the matrix is very significant. In the neutral current sector of the theory, it ensures the absence of flavour-changing neutral currents at the tree level, in accordance with experiment. In the charged-current sector, the existence of nine complex parameters which may be described using only four quantities leads to many constraints on the observables and therefore a fertile testing-ground for the theory.

The experimental form of the matrix, namely, a hierarchy of mixings, (with elements on the diagonal being close to unity and getting smaller the further away they are from the diagonal) which is reflected in the Wolfenstein parameterisation, is not predicted by any accepted theoretical principle, and although consistent with the theory, is compellingly suggestive of physics beyond the Standard Model. The hierarchical structure of the matrix has another important consequence: the unitarity condition, which is expressed by the orthonormality of the rows and columns of the matrix,

$$\sum_i |V_{i\alpha}|^2 = 1, \qquad \sum_i V_{i\alpha} V_{i\beta}^* = 0 \tag{2}$$

could, in principle, be used to determine the whole matrix once four elements are known. The constraints, being quadratic in nature, are such that even a small fractional error in one of the large (and therefore easily measured) quantities, completely dominates the smaller elements, rendering such a determination impossible. For this reason, complete determination of the matrix requires measurement of some of the smaller elements, and these are the ones which are measured in B physics. This is the origin of the strong connection between the CKM matrix and B Physics. Two of the three magnitude parameters (A and $|(\rho - i\eta)|$ in the Wolfenstein parameterisation) of the matrix are determined by direct measurements involving B-decays, whilst the best measurements of the phase are expected to come from measurements of CP violation in B decays.

The final reason for studying B physics is CP violation. At present, this phenomenon is accommodated within the Standard Model by the presence of the phase parameter in the CKM matrix. If however, there are alternative sources, (as there are in many extensions of the Standard Model) then one could expect their existence to turn-up first in an apparent violation of the unitarity of the matrix. Moreover, all CP violating observables depend on this single phase, leading to many constraints between the magnitudes of different CP-violating observables.

There are independent ways of measuring each of the four elements in the top left-hand corner of the CKM matrix. However the above discussion explains why only one of them is useful in determining the matrix as a whole. Further determination of the matrix requires measurement of the parameters involving the third generation.

The above discussion has motivated the study of the CKM matrix, and has explained the reasons for the dominance of B Physics in this study. It is of course important to test all parts of the Standard Model as accurately as possible. Other sectors of the electroweak theory have their own predictions and constraints, and much effort has been expended in performing very detailed consistency checks on these parts of the theory. One could argue that all the fundamental parameters of the Standard Model are equally important, and therefore that the most efficient use of resources is to share them equally between measurements of the various parameters. Recent advances at LEP and elsewhere in the measurement of the parameters governing the gauge sector of the electroweak theory have been truly impressive, and it is interesting to compare the current precision of the various measurements of the electroweak parameters, of which the CKM matrix parameters represent 4/17. A glance at Table 4 is surely sufficient to motivate a serious interest in improving the determination of the CKM sector of the theory. Who knows what deviations from the expectations of the model might become detectable with just one order of magnitude improvement in one of these fundamental parameters?

Table 4: Precision of various electroweak parameters.

Parameter	Value	Precision
α_{QED}	$\simeq 1/137$	3.7×10^{-7}
G_F	1.166×10^{-5} GeV$^-2$	9.0×10^{-6}
M_Z	91.187 ± 0.007 GeV$/c^2$	7.7×10^{-5}
M_W	80.40 ± 0.06 GeV$/c^2$	7.5×10^{-4}
λ	0.2196 ± 0.0023	1.0×10^{-2}
A	0.8336 ± 0.0394	4.7×10^{-2}
$(\rho^2 + \eta^2)^{1/2}$	0.41 ± 0.11	2.7×10^{-1}
η	0.28 ± 0.06	2.1×10^{-1}

2.2 Introduction to CP violation

The CP transformation is the combined operation of parity, P, in which the signs of all 3-vectors are reversed, and charge conjugation, C, in which the signs of all charges and flavour quantum numbers are reversed. The issue of CP violation is the question of whether the laws of Physics are invariant under the CP transformation, and if not, by how much is CP violated?

It was shown in 1956 that the laws of physics are not invariant under the parity transformation. In particular, it was shown that parity invariance is maximally violated by the weak interaction (it is respected by the strong and electromagetic interactions). One interpretation of this is that the expression of the laws of physics depends on whether we use a left-handed or a right-handed coordinate system. This result violates Mach's Principle, which states that the laws of physics should not depend on the kind of coordinate system used to describe them, something which would, intuitively seem very reasonable.

Soon after the discovery of parity violation, it was shown that charge conjugation invariance is also violated maximally by the weak interaction. This means that the expression of the laws of physics depends on our definition of matter and anti-matter. However, this result seemed perhaps to save Mach's Principle, because the violation of parity invariance was, apparently, exactly cancelled by the violation of charge conjugation invariance. The laws of physics for one set of coordinate axes would look the same as the laws of physics for the parity transformed set, if also, one considered anti-particles instead of particles. It looked as if it was not possible to tell the difference between the properties of matter and those of anti-matter because of the problem of defining which is left and which is right.

This situation was again turned on its head in 1964, when it was shown that CP invariance was not in fact exactly respected by the laws of nature. The K^0_L, which was thought to have a CP quantum number of -1 was found to decay into a $\pi^+\pi^-$ pair, which always has CP = + 1. It is therefore indeed possible to define which is matter, and which is anti-matter, based on some observable property of particles (other than their charge or flavour, whose signs are fixed by convention), and that, consequently, the laws of physics do look different when expressed in a left-handed coordinate system, compared with a right-handed one.

The above discussion raises a few questions about CP violation. How can CP still be violated when both P and C are maximally violated - why is the cancellation not exact? What are the observable consequences of CP violation? We will address these questions with reference to a thought experiment.

The two quadrants on the left-hand side of Fig. 7 illustrate parity violation. The top left-hand quadrant shows a polarised ^{60}Co nucleus, with its spin vector pointing downwards. When an electron is emitted simultaneously in β decay, an anti-neutrino is also emitted in the opposite direction to conserve momentum. The electron and the neutrino both have spin parallel to that of the nucleus, so that the anti-neutrino is right-handed. It is observed that the configuration shown is preferred in nature to the mirror-

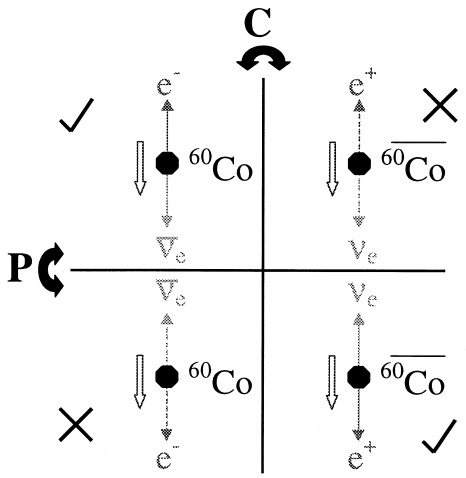

Fig. 7: Thought experiment, illustrating maximal P and C violation. Although anti-^{60}Co has never been created in the laboratory, it could be one day. The analogous experiments using pion and muon decay have been performed.

image configuration shown in the bottom left-hand quandrant. In this case, the anti-neutrino would be emitted with the opposite momentum, but the same spin direction (spin, being an axial vector, $= \mathbf{r} \times \mathbf{p}$, has the same direction in the mirror) and is therefore left-handed.

The comparison between the top left-hand quadrant, and the top right-hand quadrant of the figure illustrate C violation. All the particles in the top right-hand quadrant are the anti-particles of the ones on the left-hand side. [1] If the experiment with anti-^{60}Co were performed, the decay seen in the top right-hand quadrant of the figure would be disfavoured, compared with the one in the top left-hand corner, because of the C violation of the weak interaction. The question of maximality of P and C violation is illustrated by the fact that the disfavoured modes would be completely forbidden in the case of maximal violations, as compared with simply having reduced rates in the case of non-maximal violations. P and C asymmetries (defined as $\frac{\text{favoured}-\text{disfavoured}}{\text{favoured}+\text{disfavoured}}$) assume their maximal magnitudes of unity (at least in the limiting cases of the decay angle being exactly (anti-)parallel to the spin direction) in the case of maximal violations of these symmetries.

CP conservation is of course exemplified by the decay in the bottom right-hand corner, in which the anti-^{60}Co decays into a positron and an electron neutrino preferentially in the mirror-image case. Exact CP conservation would mean that this CP-conjugate decay would happen with exactly the same rate as that in the top left-hand figure. CP violation would be indicated, if the decay in the bottom right-hand quadrant of the figure, though preferred to the two forbidden ones, would happen at a rate different from that in the top left-hand corner.

In the case of maximal violations, there is no contradiction in the fact that both P and C symmetries are maximally violated but that the combined operation CP is still violated (the P and C violations do not exactly 'cancel out'!). The simple example of the two dis-favoured modes being exactly forbidden, and the two favoured modes having different rates, provides P and C asymmetries of maximal magnitude, unity, while still having a non-zero magnitude for the CP asymmetry.

[1] For this thought experiment, I have chosen to continue using ^{60}Co to illustrate C violation and CP conservation, by analogy with the case of parity violation, even though, of course, anti-^{60}Co has never been created in the laboratory. My aim is to illustrate the principles of C and P violation, discussed above, and use them to make some general arguments about the phenomenon of CP violation, even though the particular experiments described, are not presently possible. The analogous experiments demonstrating P and C violation have been performed using decays of pions and muons, and make interesting reading in their own right (see eg. reference [6]. They are subject to some complications, however, which are absent in the case of the experiments with anti-^{60}Co, and which make the latter a more natural example for our purposes.

Students sometimes ask: 'why is CP violation indicated by a difference between particle and anti-particle decay rates when surely C violation alone could provide that?' The thought experiment shows why this is not the case. Consider the total decay rate in the case of matter. This is obtained by integrating the differential decay rate (as a function of the direction of the electron with respect to the ^{60}Co spin direction) over all directions, including the favoured and disfavoured hemispheres. It is easy to see that such a rate will be equal between particle and anti-particle even if C is violated, as long as CP is conserved. On the other hand, as long as both C and CP are violated, then the total decay rate from the initial state to the final state is different between matter and anti-matter. This is a general rule: given that we know that C is violated in the weak interactions, *a difference in the rate for a given initial state to decay to a given final state compared to the anti-initial state to decay to the corresponding anti-final state, ie.*

$$\Gamma(\mathrm{i} \to \mathrm{f}) \neq \Gamma(\bar{\mathrm{i}} \to \bar{\mathrm{f}}) \tag{3}$$

is always a signature for CP violation. There are other signatures of CP violation, such as non-conservation of the CP quantum number in an isolated system (as in the original discovery of CP violation, in which it was found that a K_L^0 can decay to a $\pi^+\pi^-$ pair), but this is the signal which we will find most useful in the B meson system.

Note that the total decay rate for a particle to decay to all possible final states is always expected to equal the total decay rate for its anti-particle, because of invariance under the combined symmetry operation, CPT which, so far, has proved to be a good symmetry for all particle interactions.

2.3 How can CP violation arise?

CP violation is a purely quantum-mechanical phenomenon, which arises from the interference between at least two amplitudes for a given process under the following conditions:

1). there exist at least two separate amplitudes, $\mathcal{A}_i (i \geq 2)$, for the process $\mathrm{i} \to \mathrm{f}$, such that each amplitude $\mathcal{A}_i$ has two complex factors, denoted here α_i, a_i, and such that under the CP-transformation, the α_i become complex conjugated and the a_i are invariant, ie.

$$\mathcal{A}_i = \alpha_i a_i \tag{4}$$

and

$$\overline{\mathcal{A}}_i = \mathrm{CP}(\mathcal{A}_i) = \alpha_i^* a_i, \quad \text{for all} \quad i \tag{5}$$

This condition is illustrated in Fig. 8 for the case of two distinct amplitudes.

2). at least one pair of the α_i have different phases, $(\arg(\alpha_i^* \alpha_j) \neq 0, i \neq j)$

3). at least one pair of the a_i have different phases, $(\arg(a_i^* a_j) \neq 0, i \neq j)$

We can prove that these lead to CP violation in two ways. We will take the example that there are exactly two amplitudes for the process $\mathrm{i} \to \mathrm{f}$. Then, the total amplitude for the process is simply

$$\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2 \tag{6}$$

and for the CP conjugate process it is

$$\overline{\mathcal{A}} = \overline{\mathcal{A}}_1 + \overline{\mathcal{A}}_2. \tag{7}$$

The situation is shown diagramatically in the complex plane in Fig. 9. The amplitudes $\mathcal{A}_1$ and $\overline{\mathcal{A}}_1$ clearly have the same magnitude, but arguments which differ by $2\arg(\alpha_1)$. The same is true for $\mathcal{A}_2$ and $\overline{\mathcal{A}}_2$, except that they differ by the completely unrelated phase $2\arg(\alpha_2)$. So, by inspection, the magnitudes of the total amplitudes are different in the two cases:

$$|\mathcal{A}| = |\mathcal{A}_1 + \mathcal{A}_2| \neq |\overline{\mathcal{A}}| = |\overline{\mathcal{A}}_1 + \overline{\mathcal{A}}_2| \tag{8}$$

and so therefore are the corresponding rates.

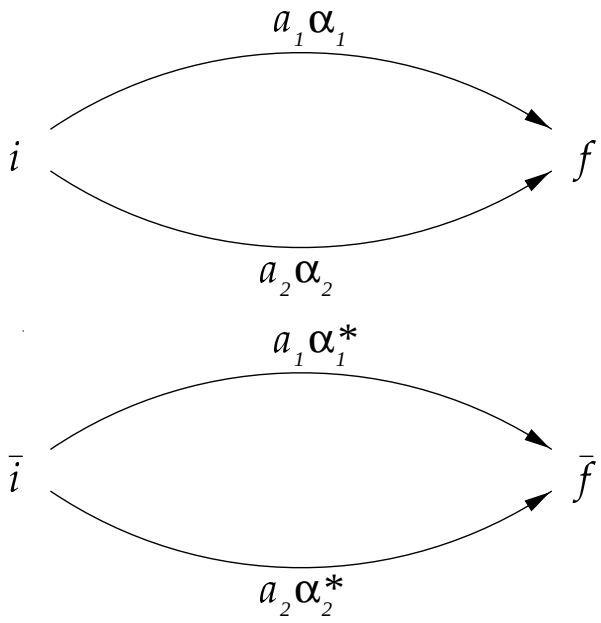

Fig. 8: CP violation requires at least two interfering amplitudes to get from initial state i to final state f. These each need to have two factors, one which is complex conjugated under CP and one which is not.

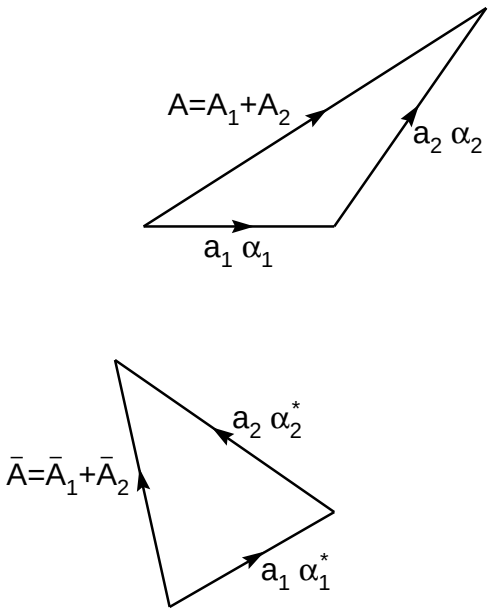

Fig. 9: Argand diagrams of the amplitude for the processes i $\rightarrow$ f (upper figure) and $\bar{\text{i}} \rightarrow \bar{\text{f}}$. The sub-amplitude $\overline{A}_1 = a_1\alpha_1^*$ differs by the phase $(2rmArg(\alpha_1))$ from A_1, and the sub-amplitude $\overline{A}_2 = a_2\alpha_2^*$ differs by the phase $(2rmArg(\alpha_2))$ from A_2.

We can be more quantitiative, by letting:

$$\mathcal{A}_i = A_i e^{i\phi_i} e^{i\delta} \tag{9}$$

where $\phi_i = \arg(\alpha_i)$, $\delta_i = \arg(a_i)$ and $A_i = |\mathcal{A}_i| = |\alpha_i||a_i|$. Then

$$|\mathcal{A}_1 + \mathcal{A}_2|^2 = (\mathcal{A}_1 + \mathcal{A}_2)(\mathcal{A}_1^* + \mathcal{A}_2^*) = A_1^2 + A_2^2 + 2A_1A_2\cos(\phi_1 + \delta_1 - \phi_2 - \delta_2) \tag{10}$$

and

$$|\overline{\mathcal{A}}_1 + \overline{\mathcal{A}}_2|^2 = A_1^2 + A_2^2 + 2A_1A_2\cos(\phi_1 - \delta_1 - \phi_2 + \delta_2) \tag{11}$$

and the difference between the rates for decay and anti-decay is

$$\begin{aligned}
\Gamma(\mathrm{i} \to \mathrm{f}) &- \Gamma(\bar{\mathrm{i}} \to \bar{\mathrm{f}}) && (12)\\
&= 2A_1A_2[\cos(\phi_1 + \delta_1 - \phi_2 - \delta_2) - \cos(\phi_1 - \delta_1 - \phi_2 + \delta_2)] && (13)\\
&= 2A_1A_2\sin(\phi_1 - \phi_2)\sin(\delta_1 - \delta_2) && (14)
\end{aligned}$$

Clearly, the difference reduces to 0 if $(\phi_1 - \phi_2) = 0$ or $(\delta_1 - \delta_2) = 0$, justifying conditions 2) and 3) for CP violation outlined above.

2.4 Mixing of neutral B mesons

I follow here quite closely the discussion in Chapter one of [2]. There are two known particle/anti-particle pairs of neutral B mesons: (B_d^0, $\overline{B}_d^0$) (composed of a b and a $\overline{d}$ quark or their anti-particles, respectively) and (B_s^0, $\overline{B}_s^0$) (made from a b and an $\overline{s}$ quark or their anti-particles). It is an experimental fact that neutral B mesons mix, which means that the B^0 and $\overline{B}^0$ transform into each other, via the weak interaction, through the second order 'Box' diagrams shown in Fig. 10. This means that, in general, a particle which is initially in one or other state (particle or anti-particle), evolves with time as an arbitrary quantum mechanical mixture of the two states. This is possible, because flavour quantum numbers are not conserved in the weak interaction.

In the rest of this section, we will not be explicit as to whether we are considering B_d^0 or B_s^0, (unless we say so) and we will use the Dirac notation to refer to the particles, $|B^0\rangle$ and $|\overline{B^0}\rangle$. We will write an arbitrary linear combination of such states as:

$$a|B^0\rangle + b|\overline{B}^0\rangle. \tag{15}$$

It can be shown that the evolution of such a state with time is governed by the time-dependent Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = H\begin{pmatrix} a \\ b \end{pmatrix} \equiv (M - \frac{i}{2}\Gamma)\begin{pmatrix} a \\ b \end{pmatrix} \tag{16}$$

for which M and Γ are two 2×2 Hermitian matrices. CPT invariance guarantees that $H_{11} = H_{22}$.

At this point, it is worth noting some of the features of the matrix H. This Hamiltonian is responsible for the propagation of the states, for their mixing, and for their decays. As such, it does not preserve the magnitude of the vector (a, b). It is therefore not Hermitian (even though, M and Γ are) and does not have real eigenvalues. It can be diagonalised in the usual way for a 2×2 matrix, and the imaginary parts of the eignevalues give the decay rates of its two eigenstates, and the real parts, their masses. In case you wonder how a Hamiltonian can be non-Hermitian, this question is resolved by noting that the two states do not form a complete basis for the process, as the final states after decay are not represented in the rows and columns of H. H is, in fact, just a small corner of the total Hamiltonian expressed in terms of all possible final states.

The off-diagonal terms of M and Γ (M_{12} and Γ_{12}, known respectively as the 'dispersive' and 'absorptive' parts of the amplitude, respectively) are given by the box diagrams shown in Fig. 10, and are,

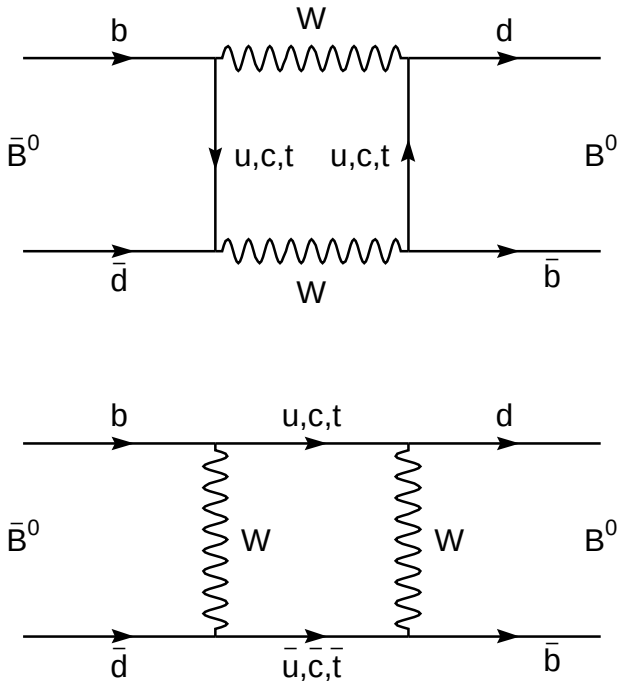

Fig. 10: Box diagrams responsible for $B^0 - \overline{B}^0$ mixing. Each is really 9 diagrams, one for each combination of u, c and t quarks as the internal lines. Interference between these diagrams is the origin of CP violation 'in the mixing'.

in general, complex quantities because of the complex CKM couplings at the vertices in the figures. M_{12} can be calculated directly from the diagrams, while Γ_{12} is calculated by cutting the quark box diagram across its centre, and assuming that the internal lines which are cut are outgoing particles in the decays. This can be justified in theoretical terms. In what follows, it will turn-out that M_{12} is more interesting, for our physics, at least in the case of B^0_d mesons.

It is also interesting to count the number of parameters involved in H. *A priori*, it appears to be a completely arbitrary complex 2×2 matrix, and would therefore have 8 parameters. However, the constraint from CPT invariance mentioned after Eq. (16) reduces this to 6. These are respectively, the (common) real and imaginary parts of the diagonal elements, and the (4) real and imaginary parts of the off-diagonal elements. How are these parameters represented in terms of observables? The observables are defined by diagonalisation of H, and they can be enumerated as the real and imaginary parts of the eigenvalues (4 parameters), and a single parameter defining the eigenvectors in terms of $|B^0\rangle$ and $|\overline{B^0}\rangle$. Where has the extra degree of freedom gone? There is a freedom in the definition of the phase of the quark fields (or equivalently, the relative phases of B^0 and $\overline{B}^0$), and the relative phase of the two off-diagonal elements of H is unobservable, thereby rendering one of the 6 parameters unobservable. The surviving observable parameter describing the eigenstates will turn out to parameterise CP violation.

The two eigenstates of H are referred to as the light, B_L, and heavy, B_H, eigenstates. These are given by

$$|B_L\rangle \;=\; p|B^0\rangle + q|\overline{B}^0\rangle, \tag{17}$$

$$|B_H\rangle \;=\; p|B^0\rangle - q|\overline{B}^0\rangle. \tag{18}$$

The complex coefficients p and q obey the normalization condition

$$|q|^2 + |p|^2 = 1. \tag{19}$$

The phase $\arg(q/p^*)$ is just the overall common phase for $|B_L\rangle$ and $|B_H\rangle$ and has no physical significance (this is yet another phase freedom, independent of the one mentioned above Eq. (17)).

The magnitudes of the masses of the two states, M_H and M_L, and their widths, Γ_H and Γ_L, will not be of interest in the rest of the story of CP violation, being dominated by the diagonal terms of H (and physically by the mass of the b quark, and strong-interaction physics). The mass difference, Δm_B, and the width difference, $\Delta\Gamma_B$, between the neutral B mesons however are of interest. They are defined as follows:

$$\Delta m_B \equiv M_H - M_L, \quad \Delta\Gamma_B \equiv \Gamma_H - \Gamma_L, \tag{20}$$

so that Δm_B is positive by definition. Finding the eigenvalues of (16), one gets

$$(\Delta m_B)^2 - \frac{1}{4}(\Delta\Gamma_B)^2 = 4(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2), \tag{21}$$

$$\Delta m_B \Delta\Gamma_B = 4\,\mathcal{R}e(M_{12}\Gamma_{12}^*). \tag{22}$$

The ratio q/p is given by

$$\frac{q}{p} = -\frac{\Delta m_B - \frac{i}{2}\Delta\Gamma_B}{2(M_{12} - \frac{i}{2}\Gamma_{12})} = -\frac{2(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}{\Delta m_B - \frac{i}{2}\Delta\Gamma_B}, \tag{23}$$

Note that its phase is unobservable, as discussed above Eq. (17). We have now discussed the 5 observables, the less interesting (for this discussion) M_B and Γ_B, and Δm_B, $\Delta\Gamma_B$, and $|q/p|$.

At this point, we need to be specific about whether we are considering B_d or B_s mesons. We will concentrate on the former. The two neutral B_d mesons are expected to have a negligible difference in lifetime,

$$\Delta\Gamma_{B_d}/\Gamma_{B_d} = \mathcal{O}(10^{-2}), \tag{24}$$

although $\Delta\Gamma_{B_d}$ has not been measured. The difference in width is produced by decay channels common to B^0 and $\overline{B}^0$. The branching ratios for such channels are at or below the level of 10^{-3}. As the various channels contribute with differing signs, it is expected that their sum will not exceed the individual level, hence $\Delta\Gamma_{B_d} \ll \Gamma_{B_d}$ is a rather safe and model-independent assumption. (For B_s^0 mesons the lifetime difference may be significant.)

On the other hand, Δm_{B_d} has been measured,

$$x_d \equiv \Delta m_{B_d}/\Gamma_{B_d} = 0.73 \pm 0.05. \tag{25}$$

From (24) and (25) we have that

$$\Delta\Gamma_B \ll \Delta m_B. \tag{26}$$

Some algebra then gives, that, to $\mathcal{O}(10^{-2})$,

$$\Delta m_B = 2|M_{12}|, \quad \Delta\Gamma_B = 2\,\mathcal{R}e(M_{12}\Gamma_{12}^*)/|M_{12}|, \tag{27}$$

$$q/p = -|M_{12}|/M_{12}. \tag{28}$$

Note that each of these quantities depends only on the off-diagonal elements of H.

The diagonalised Schrödinger equation is easily solved for each component, which in the eigenbasis have the well defined time-dependences

$$|B_L(t)\rangle = \exp\{-i(M_L - i\Gamma_L/2)t\}|B_L(0)\rangle, \tag{29}$$

$$|\overline{B}_H(t)\rangle = \exp\{-i(M_H - i\Gamma_H/2)t\}|\overline{B}_H(0)\rangle \tag{30}$$

and arbitrary states can be defined as a sum of these, with coefficients

$$a_H(t) = a_H(0)e^{-iM_H t}e^{-\frac{1}{2}\Gamma_H t}, \quad a_L(t) = a_L(0)e^{-iM_L t}e^{-\frac{1}{2}\Gamma_L t} \tag{31}$$

where $a_H(t)$ and $a_L(t)$ are the time-dependent amplitudes for the heavy and light components respectively. A state which is created at time $t = 0$ as an initially pure B^0, is denoted $|B^0_{\rm phys}\rangle$, and has $a_L(0) = a_H(0) = 1/(2p)$. Similarly an initially pure $\overline{B}^0$, $|\overline{B}^0_{\rm phys}\rangle$, has $a_L(0) = -a_H(0) = 1/(2q)$. The time evolution of these states is then given by

$$|B^0_{\rm phys}(t)\rangle = g_+(t)|B^0\rangle + (q/p)g_-(t)|\overline{B}^0\rangle, \tag{32}$$

$$|\overline{B}^0_{\rm phys}(t)\rangle = (p/q)g_-(t)|B^0\rangle + g_+(t)|\overline{B}^0\rangle, \tag{33}$$

where

$$g_+(t) = e^{-iMt}e^{-\Gamma t/2}\cos(\Delta m_B\, t/2), \tag{34}$$

$$g_-(t) = e^{-iMt}e^{-\Gamma t/2} i\sin(\Delta m_B\, t/2), \tag{35}$$

and $M = \frac{1}{2}(M_H + M_L)$, $\Gamma = \frac{1}{2}(\Gamma_H + \Gamma_L)$.

2.5 CP violation with neutral B mesons

There are three types of CP violation in B decays as follows

1. CP violation in decay matrix elements, which occurs in both charged and neutral B decays, when the amplitude for a decay and its CP conjugate process have different magnitudes. ie.

$$A(B \to f) \neq A(\overline{B} \to \overline{f}) \tag{36}$$

   where B can be B^+ or B^0. This is often called direct CP violation, and needs at least two independent amplitudes for the decay to interfere with each other. An example of this is provided by the so-called penguin diagram, an example of which can be found in the second of Fig. 12. Each penguin diagram is really a sum of three similar diagrams, each with a different charge 2/3 quark in the internal line. These diagrams have different CKM matrix elements at the vertices, which provide three different values of the phases α_i of Section 2.3 and different 'strong' phases, a_i, depending on the masses of the quarks on the internal lines. CP violation from these diagrams is difficult to calculate, as the strong phases depend on non-perturbative aspects of QCD, so it is difficult to relate what is measured to the CKM phases.

2. CP violation in the mixing, which occurs when the two neutral mass eigenstate cannot be chosen to be CP eigenstates. This is equivalent to

$$A(B^0 \to \overline{B}^0) \neq A(\overline{B}^0 \to B^0) \tag{37}$$

   This is sometimes referred to as indirect CP violation, or CP violation in the mixing. The box diagrams whose role in neutral B meson mixing was discussed in the previous section provide both the factors α_i and a_i necessary for CP violation in this case. Each box diagram is really 9 Feynman diagrams, one for each combination of charge 2/3 quarks on the internal lines. The diagrams where the two internal lines are the same charge 2/3 quarks turn out to be real, while the ones where they are different have non-zero CKM phases, and, interfering with each other as described in Section 2.3 provide CP violation. Again, the strong phases depend on the masses of the internal quarks and on non-perturbative physics. Asymmetries are expected to be small and hard to calculate.

3. CP violation in the interference between decays with and without mixing. This occurs in decays to final states that are common to B^0 and $\overline{B}^0$. (It often occurs in combination with the other two types of CP violation described above, but, there are cases when, to a good approximation, it is the only effect). This can be visualised as the interference of the process

$$B^0 \to \overline{B}^0 \to f \tag{38}$$

with

$$B^0 \to \overline{f}. \tag{39}$$

The rate obtained from the the sum of the amplitudes for these processes, differs from that for the corresponding CP conjugate process, in the way explained in Section 2.3. This type of CP violation can lead to large asymmetries in B decays. In cases where the decay matrix element is dominated by a single mode (it has negligible direct CP violation) *and* the final state is a CP-eignestate, the relationship between the observables and the CKM matrix elements turns out to be especially simple, as we will see.

2.6 CP violation in B^0 decays to CP eigenstates

Let the final state $\langle f|$ be a CP eigenstate, so that

$$|\overline{f}\rangle = \mathrm{CP}|f\rangle = \pm|f\rangle. \tag{40}$$

Then define:

$$\mathcal{A}_f = \langle f|H|B^0\rangle, \quad \overline{\mathcal{A}}_f = \langle f|H|\overline{B^0}\rangle, \tag{41}$$

and also define:

$$\lambda_f = \frac{q}{p}\frac{\overline{\mathcal{A}}_f}{\mathcal{A}_f}. \tag{42}$$

Then, the amplitude for the decay $B^0_{phys}(t) \to f$ is (using the results from Eqs. (32)-(35)):

$$\langle f|H|B^0_{phys}(t)\rangle = \mathcal{A}_f(g_+ + \lambda_f g_-). \tag{43}$$

Similarly

$$\langle f|H|\overline{B}^0_{phys}(t)\rangle = \frac{p}{q}\mathcal{A}_f(g_- + \lambda_f g_+). \tag{44}$$

As usual, the decay rate is given by the $|\text{amplitude}|^2$, i.e.

$$\Gamma(B^0_{phys}(t) \to f) = |\langle f|H|B^0_{phys}(t)\rangle|^2 \tag{45}$$

which, after some algebra gives

$$\Gamma(B^0_{phys}(t) \to f) = \mathrm{e}^{-\Gamma t}|A_f|^2[|\lambda_f|^2\cos^2(\Delta m_B t/2) + \sin^2(\Delta m_B t/2) - \mathrm{Im}(\lambda_f)\sin(\Delta m_B t)] \tag{46}$$

and more algebra gives:

$$\Gamma(\overline{B}^0_{phys}(t) \to f) = \mathrm{e}^{-\Gamma t}|\frac{p}{q}|^2|A_f|^2[\cos^2(\Delta m_B t/2) + |\lambda_f|^2\sin^2(\Delta m_B t/2) + \mathrm{Im}(\lambda_f)\sin(\Delta m_B t)] \tag{47}$$

Remember from Eq. (28) that $|\frac{p}{q}| \simeq 1$, and also that, from Eq. (40)

$$\langle\overline{f}| = \langle f|\mathrm{CP} = \pm\langle f| \tag{48}$$

so that

$$|\langle\overline{f}|H|\overline{B}^0_{phys}(t)\rangle|^2 = |\langle f|H|\overline{B}^0_{phys}(t)\rangle|^2 \tag{49}$$

Now, we define the time-dependent CP-violating asymmetry:

$$a_f(t) = \frac{\Gamma(B^0_{phys}(t) \to f) - \Gamma(\overline{B}^0_{phys}(t) \to \overline{f})}{\Gamma(B^0_{phys}(t) \to f) + \Gamma(\overline{B}^0_{phys}(t) \to \overline{f})} \tag{50}$$

which gives

$$a_f(t) = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}\cos(\Delta m_B t) - \frac{2\mathrm{Im}}{1+|\lambda_f|^2}\sin(\Delta m_B t) \tag{51}$$

CP is conserved if and only if $\lambda_f = \pm 1$.

The situation simplifies further if $|\lambda_f| = 1$. Then:

$$\begin{aligned}\Gamma(B^0_{phys}(t) \to f) &\propto \mathrm{e}^{-\Gamma t}(1 - \mathrm{Im}(\lambda_f)\sin(\Delta m_B t)) \\ \Gamma(\overline{B}^0_{phys}(t) \to \overline{f}) &\propto \mathrm{e}^{-\Gamma t}(1 + \mathrm{Im}(\lambda_f)\sin(\Delta m_B t))\end{aligned} \tag{52}$$

and

$$a_f(t) = -\mathrm{Im}(\lambda_f)\sin(\Delta m_B t). \tag{53}$$

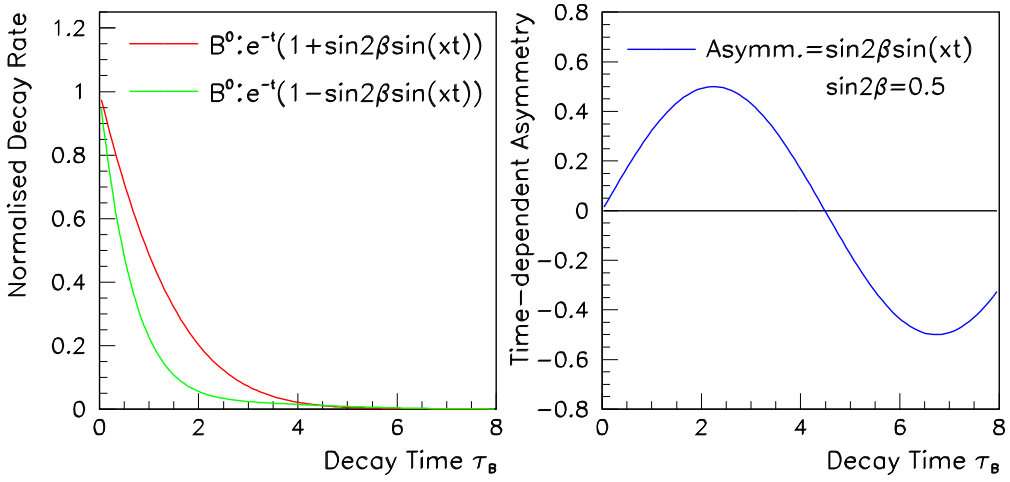

Fig. 11: Decay rate as a function of time, in units of the B^0 lifetime, for the decays $B^0 -> J/\psi K^0_S$, which is an example of Eq. (52), and the time-dependent CP asymmetry of Eq. (53). In these decays, $\mathrm{Im}(\lambda_f)$ turns out to be equal to $\sin 2\beta$, whose value is taken as 0.5 in this example.

This situation is illustrated in Fig. 11. Now, remember the definition of λ_f from Eq. (42). We also have from Eq. (28) that

$$\frac{q}{p} = -\mathrm{e}^{-i\theta} \tag{54}$$

where $\theta = \mathrm{Arg}(M_{12})$, so that the condition $|\lambda_f| = 1$ implies that

$$\left|\frac{\overline{\mathcal{A}}_f}{\mathcal{A}_f}\right| = 1 \tag{55}$$

and hence that

$$|\mathcal{A}(B^0 \to f)| = |\mathcal{A}(\overline{B}^0 \to f)| \tag{56}$$

i.e. there is negligible direct CP violation in the decay. This is therefore the condition for $|\lambda_f| = 1$ and the consequent simplifications of Eqs. (52) and (53).

The quantity $\mathrm{Im}(\lambda_f)$ depends on both the initial state, B_d^0 or B_s^0, and the final state in question. For example,

$$\left(\frac{q}{p}\right)_{B_b^0} \simeq \frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}}, \qquad \left(\frac{q}{p}\right)_{B_s^0} \simeq \frac{V_{ts}V_{tb}^*}{V_{ts}^*V_{tb}} \tag{57}$$

The most promising example, is famously $B^0 \to J/\psi\, K_S^0$ where there are, in principle, several diagrams contributing, a tree diagram, and a set of three penguin diagrams, as shown in Fig. 12

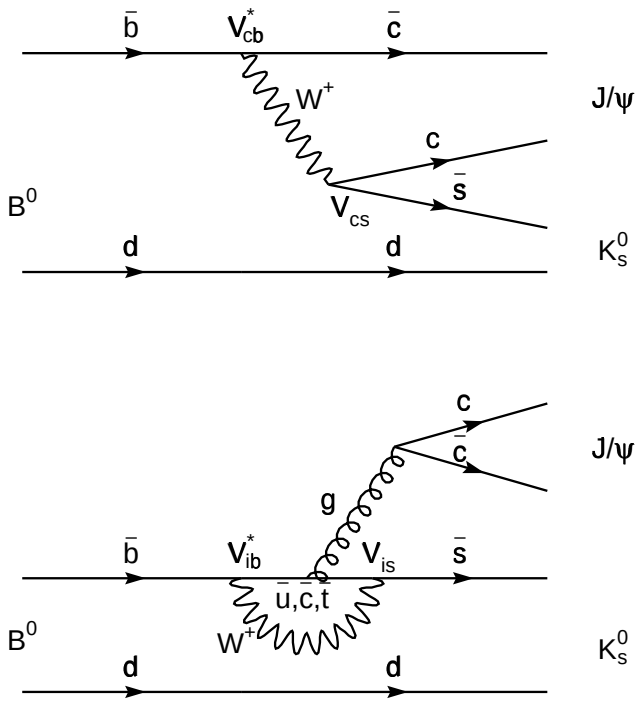

Fig. 12: Diagrams for the decay $B^0 \to J/\psi\, K_S^0$. The penguin diagram shown is really 3 diagrams, one for each internal quark line, u, c and t, which correspond to the label 'i' on the corresponding CKM matrix elements marked in the figure. The tree diagram is given by $T\times$ its CKM amplitude, $V_{cb}^*V_{cs}$. Each penguin diagram is given by $P_i\times$ its CKM amplitude $V_{ib}^*V_{is}$, where $i = u$, c or t. The quantities T and P_i correspond to the factors a_i in Section 2.3, while the CKM amplitudes, $V_{ib}^*V_{is}$ etc. correspond to the factors α_i in Section 2.3.

The total amplitude is given by:

$$\mathcal{A}_{tot} = V_{cb}^*V_{cs}(T + P_c) + V_{ub}^*V_{us}(P_u) + V_{tb}^*V_{ts}(P_t) \tag{58}$$

we can invoke the usual unitarity relation,

$$V_{tb}^*V_{ts} = -V_{cb}^*V_{cs} - V_{ub}^*V_{us} \tag{59}$$

to rewrite it as:

$$\mathcal{A}_{tot} = V_{cb}^*V_{cs}(T + P_c - P_t) + V_{ub}^*V_{us}(P_u - P_t) \tag{60}$$

and using the Wolfenstein parameterisation, we have that

$$|V_{cb}^*V_{cs}| \sim O(\lambda^2) \simeq 0.05, \qquad |V_{ub}^*V_{us}| \sim O(\lambda^4/2) \simeq 1 \times 10^{-3} \tag{61}$$

where, as usual, $\lambda = \sin\theta_c \simeq 0.22$. Hence, the quantity in Eq. (42) is, in this case, given by

$$\frac{\overline{\mathcal{A}}_f}{\mathcal{A}_f} = \frac{\overline{\mathcal{A}}_{tot}}{\mathcal{A}_{tot}} = \frac{V_{cb}^* V_{cs}}{V_{cb} V_{cs}^*} \tag{62}$$

Finally, and exceptionally in the case of a K_s^0 in the final state, we have an extra phase which comes from the $K_s^0 - \overline{K}_s^0$ mixing:

$$\left(\frac{p}{q}\right)_{K^0} \simeq \frac{V_{cb}^* V_{cs}}{V_{cb} V_{cs}^*} \tag{63}$$

Putting all this together gives:

$$\lambda_f = \left(\frac{V_{td}^* V_{tb}}{V_{td} V_{tb}^*}\right)\left(\frac{V_{cb}^* V_{cs}}{V_{cb} V_{cs}^*}\right)\left(\frac{V_{cs}^* V_{cd}}{V_{cs} V_{cd}^*}\right) = \mathrm{e}^{i2\beta} \tag{64}$$

$$\Rightarrow \mathrm{Im}(\lambda_f) = \sin(2\beta). \tag{65}$$

This result is typical of the case of a final CP eigenstate with the decay dominated by a single amplitude. In such cases, $\mathrm{Im}(\lambda_f)$ is always given by the sine of twice an angle in the Unitarity triangle.

2.7 The generic CP asymmetry measurement

In order to observe a time-dependent CP asymmetry, it is necessary to reconstruct both a sample of B^0 decays to the final state in question, and one of $\overline{B}^0$. Under the assumption that they are produced in equal quantities, and detected with equal efficiencies, the asymmetry is then

$$a_f(t) = \frac{N(t) - \overline{N}(t)}{N(t) + \overline{N}(t)} \tag{66}$$

where $N(t)$ and $\overline{N}(t)$ are respectively the number of initial B^0 and $\overline{B}^0$ particles observed to decay to the final state, as a function of the proper time, t. In the e^+e^- case, this proper time is the proper time difference between the decays of the two associatively produced B mesons in each event. In the hadronic case, it is the proper time between the production and decay of the B^0 which decays to the CP mode in question.

In order to begin to make a significant measurement of CP violation, a few hundred reconstructed events are needed in each sample. Time-dependent measurements require vertex detectors, so that the decay vertices of the neutral B-mesons can be observed.

In the cases which are theoretically cleanest, namely final CP-eigenstates, the final state is the same for both initial states and therefore does not allow us to distinguish between them. It is therefore necessary to tag the initial states in some other way. In both e^+e^- collisions and at hadron machines, neutral B-mesons are produced in association with a partner of the opposite beauty (but not necessarily the anti-particle at hadron machines). If this can be identified, and the sign of its beauty detected, then the beauty of the decaying particle is tagged. This is done using the lepton sign in semi-leptonic decays of the associated B-meson, and usually also using the charge of kaonic B-decay products. The observed asymmetry is then defined as

$$a_{f,obs}(t) = \frac{N^+(t) - N^-(t)}{N^+(t) + N^-(t)} \tag{67}$$

where the superscript denotes the charge of the tagging particle for a given decay at proper time, t.

The tagging is not perfectly efficient, as branching ratios and acceptances for the tagging particle have to be taken into account. It is furthermore not perfectly accurate, and several sources of wrong-sign tags are present. These include:

- $B^0 - \overline{B}^0$ mixing of the tagging B-meson.
- Semi-leptonic decays of charmed decay products of the tagging B-meson, which have the wrong sign.
- Leptonic decays of other hadrons produced in the same event, whether beauty, charm or light flavours.
- Fake leptons and wrong charge measurements of the tagging lepton.
- Wrong-sign kaon tags: although a given B-meson has a preferred charge for its kaonic products, either sign of kaon can result, albeit with different probabilities.

These mis-tags lead to an effective dilution of the asymmetry. Defining the dilution factor as

$$D = \frac{N_g - N_b}{N_g + N_b}, \tag{68}$$

where N_g is the number of good tags and N_b is the number of bad tags, it is easy to see that the observed asymmetry defined in Eq. (67) is given by

$$a_{f,obs}(t) = D \times a_f(t) = D \sin 2\phi_f \sin(\Delta m t) \tag{69}$$

where ϕ_f is the relevant angle of the unitarity triangle, as listed in Section 1.2. D is clearly less than one and reduces the effect, leading to an increase in the number of events needed to observe the asymmetry.

Having observed the time-dependent asymmetry, the sinusoidal form may be fitted, and its amplitude $D \sin 2\phi_f$ extracted. A correction must be made for D, which can either be measured using a CP-conserving channel, or estimated using Monte Carlo, although the former is clearly more desirable. The uncertainty in D is one of the major systematic errors in the experimental determination of the angles of the unitarity triangle.

3. CONCLUSION

The expected sensitivity of the BABAR experiment to the CKM parameters ρ and η with 90 fb^{-1} of data collected, i.e. three years' running at design luminosity is shown in Fig. 13. The two solutions for β correspond to a single measurement of $\sin 2\beta$. The 'longitudinal' extent of the grey (yellow) areas corresponds to the sensitivity to the CP angle α, and it statistics-limited in this case.

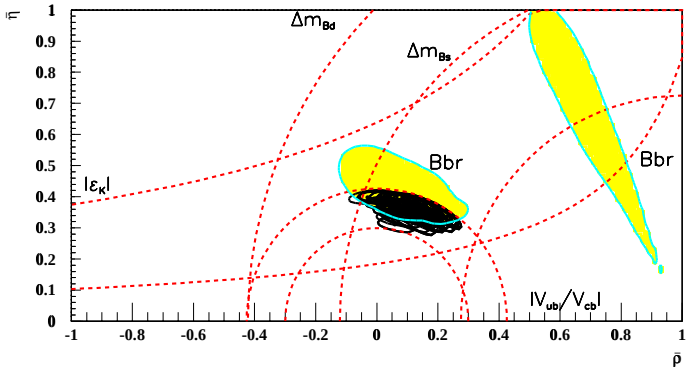

Fig. 13: Expected sensitivity to the CKM parameters ρ and η at BABAR with 90 fb^{-1} of data collected. The dashed lines indicate the expected 95% CL constraints from various measurements at that time, which do not involve CP-violation measurements, and the envelope of the black ellipses indicates the combined sensitivity of all those measurements. The two grey-shaded (or yellow) areas indicate the sensitivity using CP-violation measurements only. In this example, we have have chosen a set of measurements which are consistent. They may of course turn-out in reality to be inconsistent, in which case, the areas would not overlap, and a deviation from the standard model would have been discovered.

It is clear that there is a bright future ahead for B factories, in the next few years, and that the stage is set for several exciting tests of the standard model.

ACKNOWLEDGEMENTS

I would like to thank the summer school organizers for inviting me to give these lectures. I also thank Theresa Champion for helping me with some of the figures, and my collaborator, Bill Scott, for continuing collaboration on issues related to CP violation.

References

[1] D. Wyler, these proceedings.

[2] The BABAR Collaboration, 'The BABAR Physics Book', eds. P.F. Harrison and H. Quinn, SLAC-R-504, 1998. http://www.slac.stanford.edu/pubs/slacreports/slac-r-504.html

[3] H. Quinn in 'CP Violation in B Decay - Standard Model Predictions', Review of Particle Physics, Eur. Phys. J. **C3** (1998) 558. http://pdg.lbl.gov/

[4] Y. Nir in Proceedings of the 1998 European School of High-Energy Physics, CERN 99-04 (1999) 99. http://preprints.cern.ch/cernrep/1999/99-04/99-04.html

[5] J.D. Richman in 'Heavy Quark Physics and CP Violation' Lectures at the 1997 Nato Advanced Study Institute, Les Houches, France. http://www.hep.ucsb.edu/papers/driver_houches12.ps

[6] E.D. Commins, 'Weak Interactions', McGraw-Hill, 1973.